Chapter 1
Random Processes

Prepared by Prof

Mahmoud Ahmed Attia Ali

Department of Electronics and Communications
Faculty of Engineering
Tanta University
Egypt

TIP
For additional advice see
Dale Carnegie Training®
Presentation Guidelines

Probability Theory and Stochastic Processes

A basic tool designing of digital communication systems including

- Modeling sources that generate the information.
- Digitization of the source output.
- Characterizing of various communication channels.
- Designing of the receiver structure.
- Evaluation the performance of the whole system.

Probability Theory

Concepts of Probability

Deals with averages of mass phenomena occurring sequentially or simultaneously

- Electron emission,
- Telephone calls,
- Radar detection,
- Quality control,
- System failure,
- Birth and death rates,

The purpose is to predict such averages in terms of probabilities of events.

Probability of an Event

If an experiment is performed N times and the event A occurs N_a times, then the probability of A is:

$$P(A) = \frac{N_a}{N}$$

- N_a: *Number of outcomes in favor to event* A.
- N : Number of all possible outcomes.

Union and Intersection

Union $A+B$ of two events is the event that occurs when A or B or both occur:

Intersection, or joint $A.B$ is the event that occurs when both events A and B occur.

- If A and B interleaved each other, a union and joint probabilities are interrelated by:

$$P(A+B) = P(A) + P(B) - P(A.B)$$

- If A and B are independent, a union and joint probabilities are unrelated so that:

$$P(A+B)=P(A)+P(B)$$

Conditional Probability

If outcomes of two events A and B form their joint $A.B$, the conditional probability of event A given event B is defined as:

$$P(A/B) = \frac{P(A.B)}{P(B)}$$

If events A and B are independent, then there is no joint occurrence, so that:

$$P(A/B) = P(A)$$

So, joint of two independent events equals the product of their individual probabilities:

$$P(A.B) = P(A).\ P(B)$$

Binomial Distribution

Assume p is the probability of occurrence of an event A. So, probability of non-occurrence of such event will be termed as $q = 1 - p$.

After repeating the experiment N times, the probability that A is observed exactly k times out of N trials, is a Binomial distribution:

$$P(A\ occurs\ k\ times) = \frac{N!}{k!\,N-k!}\,p^k\,(1-p)^{N-k}$$

Binomial Applications

It involves any experiment for which there are only two possible outcomes on any trial.

- Hitting or missing the target in artillery,
- **Passing or failing an exam,**
- Receiving `0' or `1' in a digital bit stream transmission, and
- Occurrence or non-occurrence of any event.

Random Variables

Random Variables

A random variable is a number assigned to every outcome of an experiment. It could be:

- Voltage of a random source,
- Phase of a random signal,
- Power of a received signal, or any other.

It is characterized by three basic functions

- Cumulative distribution function cdf.
- Probability density function pdf.
- Mass function.

Cumulative Distribution Function

Probability that: Random variable X takes on a value equal to or less than x in a trail.

$$F_X(x) = P(X \le x)$$

The discrete random variable by flipping a fair coin has cdf in Fig1a, with 2 steps.

The random variable by tossing a fair die in Fig1b of 6 jumps, one at $x = 1, 2, \ldots, 6$.

Examples on

Cumulative Distribution Functions

There are two jumps in $F_X(x)$,

- One at x = -1 and
- One at x = 1.

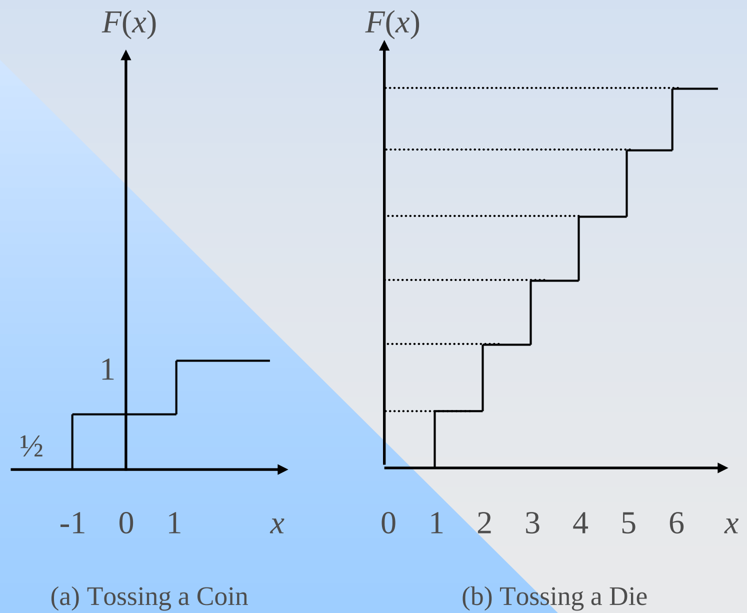


Fig.1.1 Examples of Discrete CDF

Properties of cdf

$$F_X(-\infty) = 0$$

$$F_X(\infty) = 1$$

$$0 \le F_X(x) \le 1$$

$$F_X(x_1) \le F_X(x_2), \qquad \text{if} \;\; x_1 \le x_2$$

$$P(x_1 < X \le x_2) = F_X(x_2) - F_X(x_1)$$

Probability Density Function

pdf is the derivative of the cumulative distribution function cdf:

$$f_X(x) = \frac{d}{dx} F_X(x)$$

So, cdf, can be given in terms of the density function, pdf, by integration:

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\, dx$$

Properties of pdf

$$f_x(x) > 0 \text{ for all } x$$

$$\int_{-\infty}^{\infty} f_x(x)\,dx = 1$$

$$F_X(x) = \int_{-\infty}^{x} f_x(y)\,dy$$

$$P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f_x(x)\,dx$$

Mass Function

Concerned with discrete variables as:

$$p_x(x) = P(X = x)$$

For random variable of discrete values x_i

$$p_i = P(X = x_i)$$

pdf can be evaluated in terms of the mass functions as follows:

$$f_x(x) = \sum_i p_i \, \delta(x - x_i)$$

Two Random Variables

Two Random Variables

Over finite interval $x_1 \le X \le x_2$, and $y_1 \le Y \le y_2$:

$$P(x_1 \le X \le x_2, y_1 \le Y \le y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} f_{xy}(x, y)\, \mathrm{dx}\, \mathrm{dy}$$

So, cdf is:

$$F_{xy}(x, y) = P(X \le x, Y \le y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{xy}(x, y)\, \mathrm{dx}\, \mathrm{dy}$$

For x quite independent on y:

$$F_x(x) = P(X \le x, -\infty \le Y \le \infty) = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f_{xy}(x, y)\, \mathrm{dx}\, \mathrm{dy}$$

$$f_x(x) = \frac{d}{dx}[F_x(x)] = \int_{-\infty}^{\infty} f_{xy}(x, y)\, \mathrm{dy}$$

Independent Random Variables

If variables *X* and *Y*, are independent:

$$P(x \leq X \leq x + dx, y \leq Y \leq y + dy) = [f_x(x)dx][f_y(y)dy]$$

$$P(x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2) = \left[\int_{x_1}^{x_2} f_x(x)dx\right]\left[\int_{y_1}^{y_2} f_y(y)dy\right]$$

$$f_{xy}(x, y) = f_x(x)\, f_y(y)$$

Statistical Averages

Average value is the expectation of the random variable *X*, and is given by:

$$\overline{X} = E[X] = m = \sum_i x_i P(x_i)$$

For a continuous random variable:

$$\overline{X} = E[X] = m = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

For a function $g(x)$ of the variable *X* is:

$$\overline{g(x)} = E[g(x)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx$$

n^{th} Moment

If the random variable X is raised to a power n, the average value of X^n is referred to as the n^{th} moment of the random variable X given by:

$$\overline{X^n} = E[X^n] = \int_{-\infty}^{\infty} X^n f(x)\, dx$$

Variance

Variance σ^2 is a measure of the width of the probability density function:

Is equivalent to the average value of the second moment $(X\text{-}m)^2$ as:

$$\sigma^2 = E[(X-m)^2] = \int_{-\infty}^{\infty} (x-m)^2 f(x)\,\mathrm{dx}$$

$$\sigma^2 = E[X^2 - 2mX + m^2]$$
$$= E[X^2] - 2mE[X] + m^2 = E[X^2] - 2m^2 + m^2$$
$$= E[X^2] - m^2 = E[X^2]_{if\ m=0}$$

Gaussian Distribution

Many natural events are characterized by Gaussian density such as thermal noise. :

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-m)^2/2\sigma^2}$$

$$m = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi\sigma^2}} e^{-(x-m)^2/2\sigma^2} dx$$

$$\sigma^2 = \int_{-\infty}^{\infty} \frac{(x-m)^2}{\sqrt{2\pi\sigma^2}} e^{-(x-m)^2/2\sigma^2} dx$$

$$\int_{-\infty}^{\infty} f(x) = 1$$

Error Function

Cumulative function cdf of Gaussian density $f(x)$ with zero mean is:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-x^2/2\sigma^2} dx$$

This integral is not easy and is available in tables and is termed error function written as:

$$\mathrm{erf}\, u = \frac{2}{\sqrt{\pi}} \int_0^u e^{-u^2} du$$

$$\mathrm{erf}(0) = 0, \qquad \mathrm{erf}(\infty) = 1$$

Complementary Error Function

Complementary error function erfc u defined:

$$\operatorname{erfc} u = 1 - \operatorname{erf} u = \frac{2}{\sqrt{\pi}} \int_{u}^{\infty} e^{-u^2}\,\mathrm{du}$$

One could prove that cumulative distribution of Gaussian process is given as:

$$F(x) = \frac{1}{2} \operatorname{erfc}\left(\frac{|x|}{\sqrt{2}\sigma}\right)$$

Density of $y = g(x)$

- Given a random variable X where its associated density function is $f_X(x)$
- What is the density function of random variable $Y = g(x)$ where $g(x)$ is some function of X?

Density of $y = g(x)$

Solve the equation:

$$Y = g(x).$$

Assume it has x_n real roots,

The density function will be:

$$f_y(y) = \frac{f_x(x_1)}{g'(x_1)} + \frac{f_x(x_2)}{g'(x_2)} + \cdots + \frac{f_x(x_n)}{g'(x_n)}$$

Examples of $g(x)$

1) If $g(x)$ is a linear function, $Y = a\,X + b$:

- It has one solution $x = (y - b/a)$ and
- The derivative is $g'(x) = a$, hence:

$$f_y(y) = \frac{f_x(x_1)}{g'(x_1)} = \frac{1}{a} f_x\left(\frac{y-b}{a}\right)$$

2) If $g(x)$ is the inverse, $Y = 1/X$:

- It has one solution $x = 1/y$,
- The derivative is $g'(x) = -\frac{1}{x^2}$ corresponds to y^2:

$$f_y(y) = \frac{f_x(x_1)}{g'(x_1)} = \frac{1}{y^2} f_x\left(\frac{1}{y}\right)$$

$g(x)$ is Sinusoidal

If $g(x)$ is sinusoidal, $Y = a\sin(X + \emptyset)$

It has n solutions $f_x(x_n)$

The derivatives is: $g'(x) = a\cos(x + \emptyset)$

$$\therefore y^2 + [g'(x)]^2 = a^2\sin^2(x + \emptyset) + a^2\cos^2(x + \emptyset) = a^2$$

$$\therefore [g'(x)]^2 = a^2 - y^2 \qquad or \qquad \therefore g'(x) = \sqrt{a^2 - y^2}$$

Hence, the density function of Y is given by:

$$f_y(y) = \frac{1}{\sqrt{a^2 - y^2}} \sum_{n=-\infty}^{\infty} f_x(x_n)$$

Central Limit Theory

Density of sum of N independent Gaussian random variables approach Gaussian density as N increases.

Its mean is the same mean of N independent random variables, whereas standard deviation is $\sigma/\sqrt{N}$

This theorem is even applied when individual random variables are not Gaussian.

It is also applies in certain special cases even when individual random variables are not independent.

Random Process

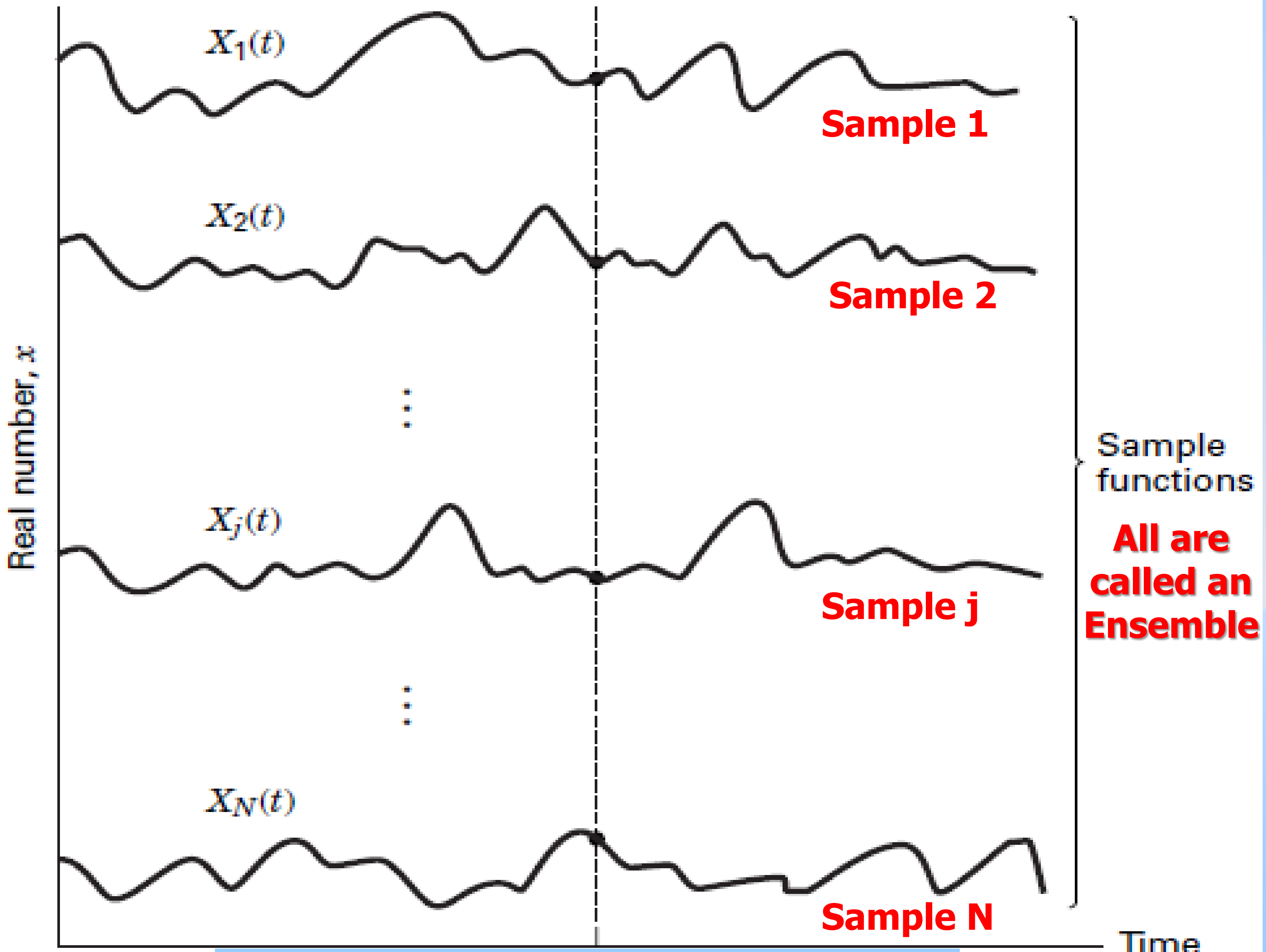
Random Process

A random process $X(A,t)$ is a function of two variables: an event A and time t.

Fig illustrates random process with N sample functions of time $\{X_j(t)\}$each as the output of different noise generator.

- For a specific event A_j, we have single time function $X(A_j,t) = X_j(t)$ (i.e. single sample function).
- For a specific time t_k, $X(A,t_k)$ is a random variable $X(t_k)$ whose value depends on the events.
- For specific event A_j and specific time $t = t_k$, $X(A_j,t_k)$ is simply a number.

The totality of all sample functions is called an ensemble.



Random Noise Process

How to Describe

Statistical Properties

A random process whose distribution functions are continuous can be described statistically with (pdf).

- But generally, pdf of a random process will be different for different times.
- Also in most cases it is not practical to get empirically the probability distribution of a random process.

However, a partial description consisting of the mean and autocorrelation function are often adequate for the needs of communication systems.

We define mean & autocorrelation of random process $X(t)$ as:

Mean & Autocorrelation

We define the **mean** of the random process $X(t)$ as:

$$E[X(t_k)] = \int_{-\infty}^{\infty} x\, p_{X_k}(x)\ \mathrm{dx} = m_x(t_k)$$

$X(t_k)$; the random variable obtained by observing the random process at time t_k

$p_{X_k}(x)$; pdf of $X(t_k)$ over the ensemble of events at t_k.

Autocorrelation of random process $X(t)$ is a function of two times t_1 and t_2: (measure of the degree to which 2 time samples of the same random process are related)

$$R_x(t_1, t_2) = E[X(t_1)\, X(t_2)]$$

$X(t_1)$ and $X(t_2)$; are the random variables obtained by observing $X(t)$ at times t_1 and t_2.

Stationary Process

Stationary Process

Random process $X(t)$ is said to be strict-sense stationary if none of its statistics are affected by a shift in time.

And, is said to be wide-sense stationary (WSS) if two of its statistics (mean and autocorrelation) do not vary with a shift in time. Thus, a process is WSS if

$$E[X(t)] = m_x = constant$$

$$R_X(t_1, t_2) = R_X(t_1 - t_2)$$

Strict-sense implies wide-sense, but not vice versa.

Most results in communication theory are predicated on random signals and noise being wide-sense stationary.

In practice, it is not necessary for a random process to be stationary for all time but only for interval of interest.

Autocorrelation of WSS

Variance; a measure of randomness for random variables, Autocorrelation; a measure for random processes.

For WSS process, autocorrelation is a function of time difference $\tau = t_1 - t_2$;

$$R_X(\tau) = E[X(t)X(t+\tau)] \quad \text{for} \quad -\infty < t < \infty$$

For zero mean WSS process, $R_X(\tau)$ indicates the extent to which random values of process separated by τ seconds are statistically correlated.

It gives an idea of frequency response that is associated with a random process.

Autocorrelation and Response

It gives an idea of frequency response that is associated with a random process.

If it changes slowly as τ increases, it indicates that, sample values of $X(t)$ taken at t_1 and $t_1 + \tau$ are nearly the same. Thus, we would expect a frequency domain of $X(t)$ to contain low frequencies.

Or, if it decreases rapidly as τ is increased. $X(t)$ expected to change rapidly with time and thereby contain mostly high frequencies.

Properties of autocorrelation of a real-valued wide-sense stationary process are in Table.1.1.

Table.1.1: Properties of Autocorrelation of a Real-valued Wide-sense Stationary Process

n	Property	Meaning
1	$R_X(\tau) = R_X(-\tau)$	Symmetrical in τ about zero
2	$R_X(\tau) \leq R_X(0)$ all τ	maximum value occurs at the origin
3	$R_X(\tau) \leftrightarrow G_X(f)$	Autocorrelation and power spectral density form a Fourier transform pair
4	$R_X(0) = E\{X^2(t)\}$	Value at the origin is equal to the average power of the signal

Ergodic Process

Time & Ensemble Average

With ensemble averaging, we have to average across all the sample functions of the process and would need to have complete knowledge of the first and second-order joint pdfs. Such knowledge is generally not available.

When random process is **ergodic**, its time averages equal to its ensemble averages, and the statistical properties of process can be determined by time averaging over a single sample function of the process.

To be ergodic, it must be stationary in strict sense.

Ergodic process is stationary, but, stationary process is not necessarily ergodic.

Ergodic Process

For most communication, we are satisfied to meet conditions of wide-sense stationarity, we are interested only in the mean and autocorrelation.

Random process is ergodic in mean and autocorrelation if:

$$m_X = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t)\, dt$$

$$R_X(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t)\, X(t+\tau) dt$$

Reasonable assumption in most communication signals (in absence of transient) is that the random waveforms are ergodic in the mean and the autocorrelation function.

Power Spectral Density

Power Spectral Density

Power spectral density *psd* is very useful in communication, it describes distribution of a signal 's power in the frequency domain.

It enables us to evaluate the signal power that will pass through a network having known frequency characteristics.

We summarize the principal features of *psd* as follows:

Features of *psd*

- $G_X(f) \geq 0$ and is always real valued
- $G_X(f) = G_X(-f)$ for $X(t)$ real valued
- $G_X(f) \leftrightarrow R_X(\tau)$
- Relationship between average normalized power and *psd*:

$$P_X = \int_{-\infty}^{\infty} G_X(f)\, df$$

Noise

Noise

- **Define:** Unwanted signals that disturb transmission and processing of signals in communication systems.
- **Types:** There are many sources of noise:

1-External: atmospheric, galactic, man-made noise

2-Internal to the system:

Shot noise, arises because of the discrete nature of current flow in electronic devices.

Thermal noise, which is due to the random motion of electrons in a conductor.

- **Analysis:** Analysis is usually based on a source of noise called **white-noise**.

Stationary Noise

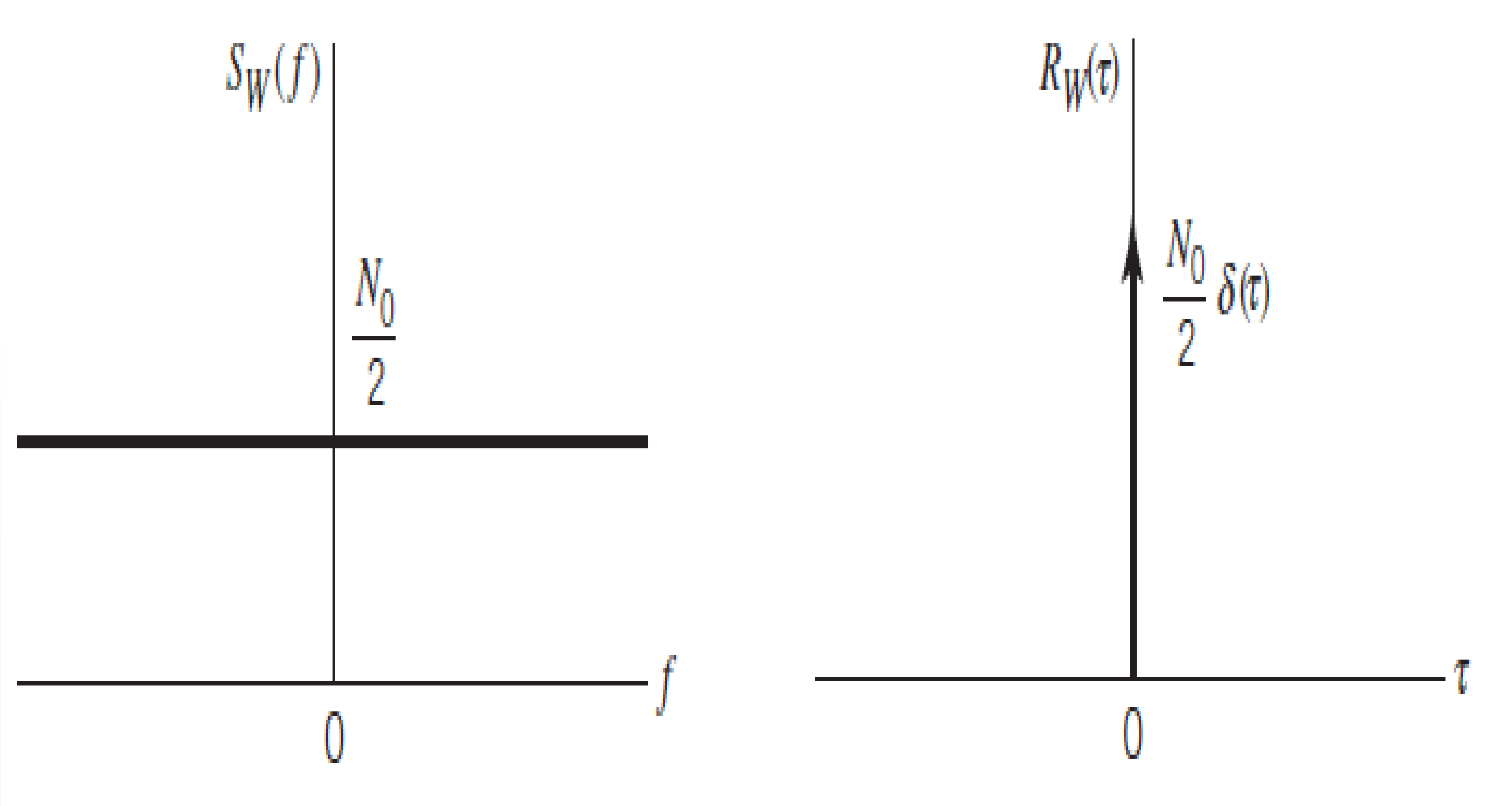
White noise, denoted by $W(t)$, is a **stationary process** whose power spectral density $S_W(f)$ has a constant value across the entire frequency interval.

$$S_W(f) = \frac{N_o}{2}$$

Autocorrelation $R_W(\tau)$ is the inverse Fourier transform of power spectral density (Wiener–Khintchine):

$$R_W(\tau) = \frac{N_o}{2}\,\delta(\tau)$$

Since $R_W(\tau)$ is zero for $\tau \neq 0$, so any two different samples of white noise are uncorrelated no matter how closely together in time those two samples are taken.



Power Spectral Density (b) Autocorrelation Function

Characteristics of White Noise

Characteristic of Noise

White noise is also **Gaussian**, so the two samples are statistically independent. Thus, white Gaussian noise represents the ultimate in "**randomness**".

Utility of **white-noise** process in the noise analysis of communication systems is **parallel to** the use of **delta function** in the analysis of linear systems.

As long as **bandwidth** of noise process at the input of a system is larger than the bandwidth of the system itself, then we may model noise process as **white noise**.

Ideal Low Pass Filtered Noise

Ideal LP Filtered Noise

If white Gaussian noise of zero mean and $N_o/2$ power spectral density is applied to ideal LPF of band B, output

$$S_N(f) = \begin{cases} \dfrac{N_0}{2}, & -B < f < B \\ 0, & |f| > B \end{cases}$$

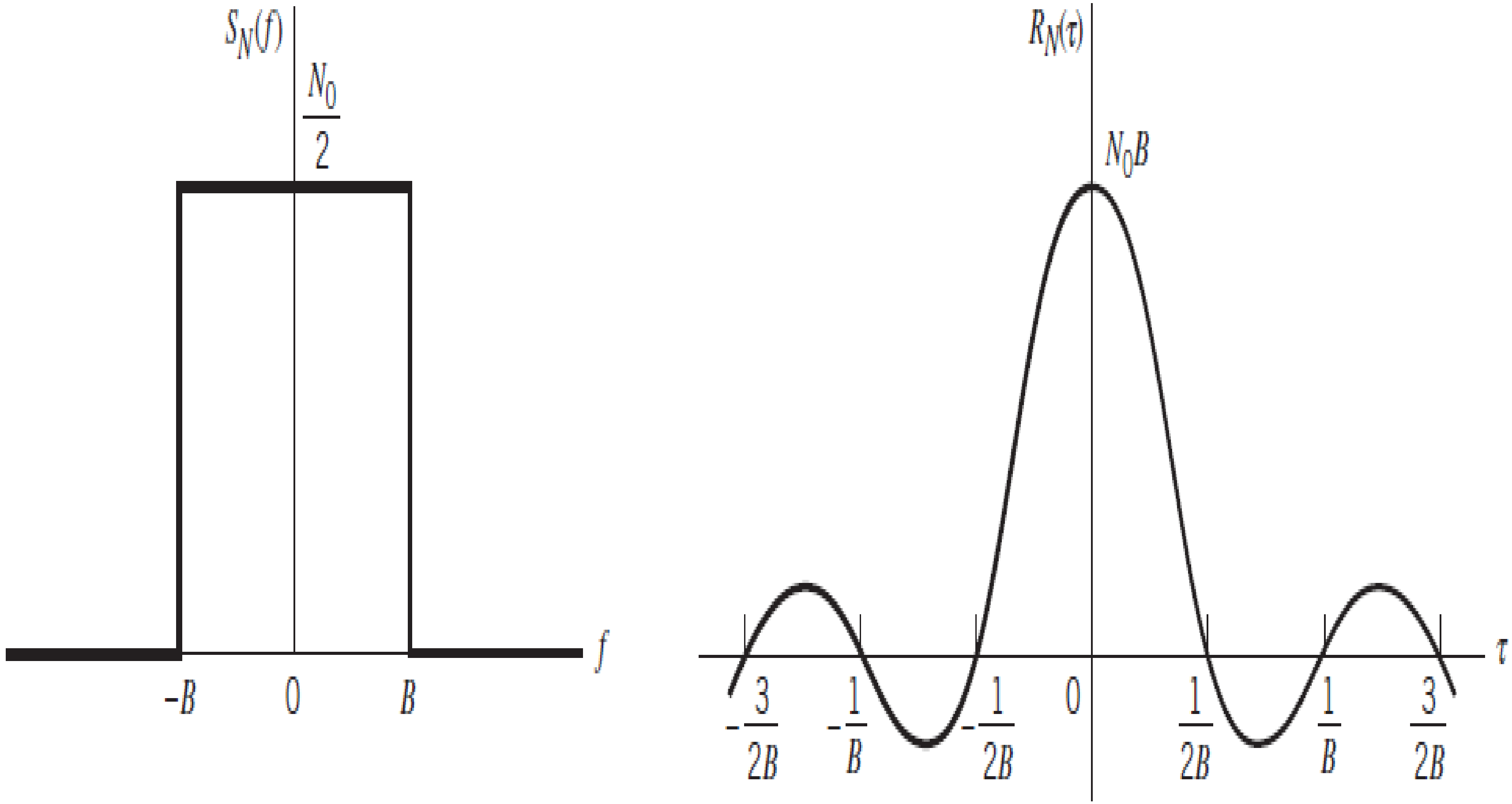
Autocorrelation function is inverse Fourier transform:

$$R_N(\tau) = \int_{-B}^{B} \frac{N_o}{2} e^{j2\pi f\tau}\, df = N_o B \operatorname{sinc}(2B\tau)$$

So, If input noise is Gaussian, band-limited noise $N(t)$ at filter output is also Gaussian.

If $N(t)$ is sampled at $2B$, resulting noise samples are uncorrelated, Gaussian, and statistically independent.

Ideal LP Filter



(a) Power Spectral Density (b) Autocorrelation Function

Characteristics of Low Pass Filtered White Noise

Noise Correlation with Sinusoid

Correlation of Noise with sinusiod

Consider output of correlator with white Gaussian noise $w(t)$ and sinusoidal wave at its two inputs;

$$w'(t) = \sqrt{\frac{2}{T}} \int_0^T w(t) \cos(2\pi f_c t)\, dt$$

If $w(t)$ of zero mean, output $w'(t)$ has zero mean too.

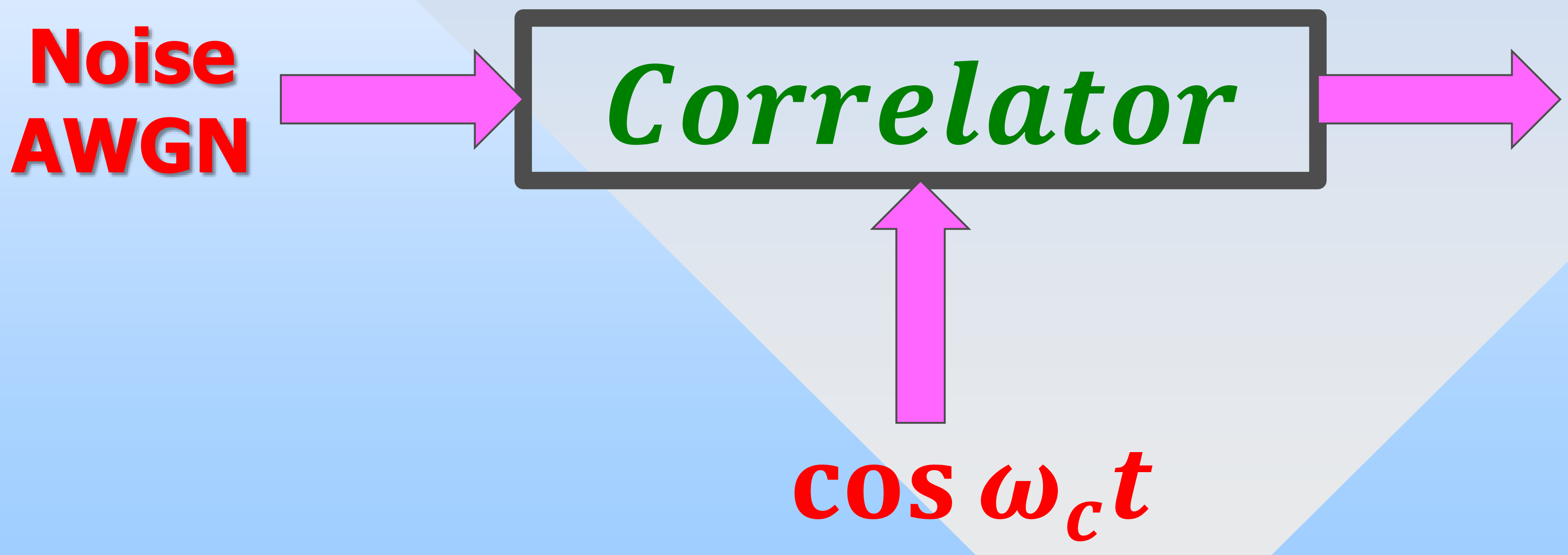
Variance of correlator output is therefore defined by:

$$\sigma_{W'}^2 = E\left[\frac{2}{T} \int_0^T \int_0^T w(t_1) cos(2\pi f_c t_1) w(t_2) cos(2\pi f_c t_2)\, dt_1 dt_2\right]$$

$$\sigma_{W'}^2 = \frac{N_o}{2}$$

Assume the frequency f_c of sinusoid is integer multiple of the reciprocal of T for mathematical convenience.

Correlator



Narrowband Noise

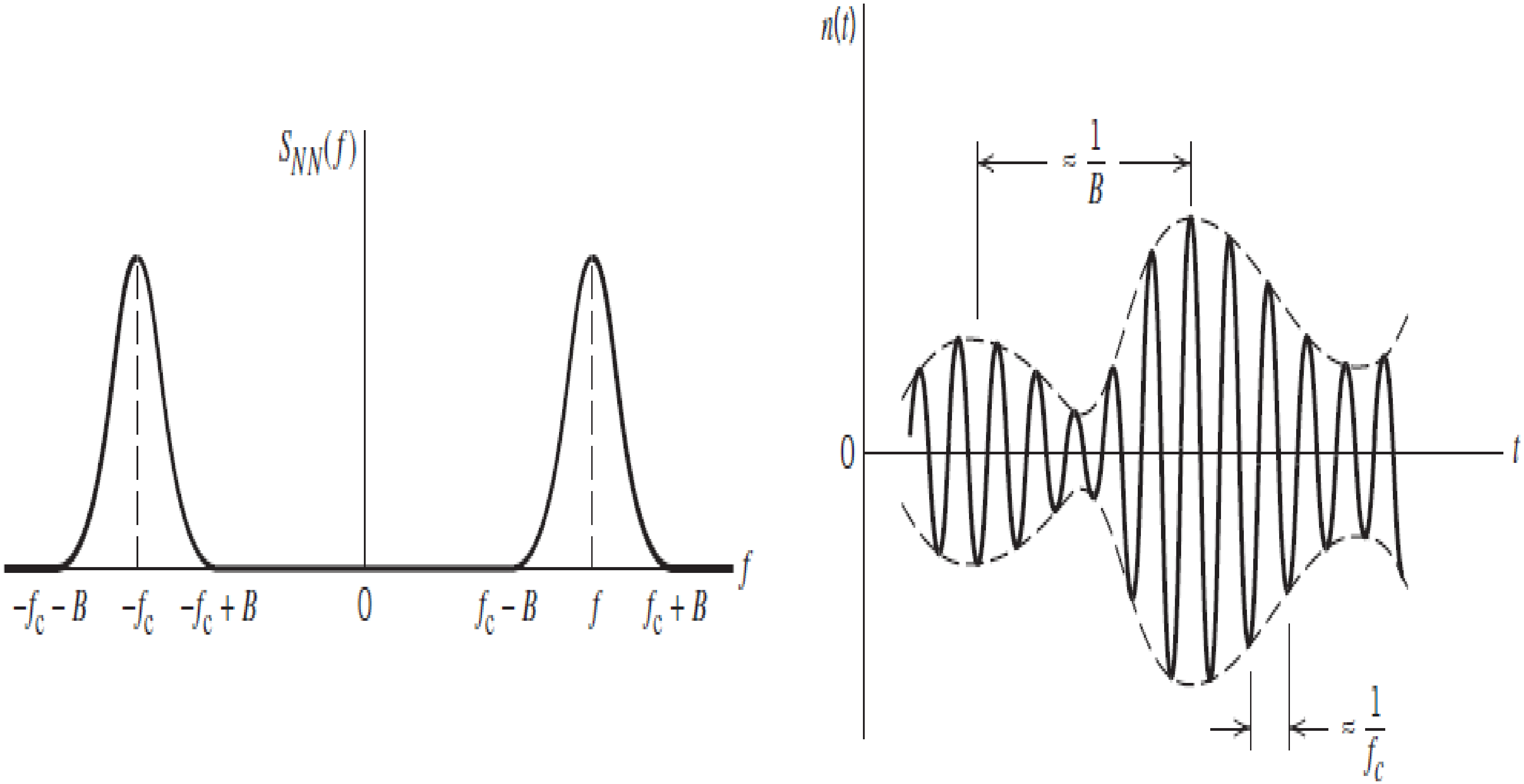
Narrowband Noise

Receiver includes **preprocessing** for received signal. It takes a form of narrowband filter of enough band to pass modulated component of received signal undistorted, to limit the effect of channel noise passing in the receiver.

The noise process appearing at the output of such a filter is called narrowband noise.

If spectral components of narrowband noise centered about midband frequency $\pm f_c$, we find that a sample function $n(t)$ of such a process appears somewhat similar to a sine wave of frequency f_c.

The sample function $n(t)$ may, then, change slowly in both amplitude and phase, as in Fig.



(a) Power Spectral Density **(b) Sample Function**

Narrowband Noise

Narowband Noise

Consider $n(t)$ produced at narrowband filter output in response to the sample *w(t)* of white Gaussian noise process of zero mean and unit power spectral density of the respective processes $N(t)$ and $W(t)$.

We express power spectral density of $N(t)$ in terms of the transfer function of filter $H(f)$ as:

$$S_N(f) = [H(f)]^2$$

So, any narrowband noise encountered in practice may be modeled by applying a white-noise to a suitable filter in the manner described as above.

In phase and Quadrature Components

Inphase and Quadrature Components of Noise

Oftenly, it is convenient to represent the narrowband noise $n(t)$ in terms of its in-phase component $n_I(t)$ and quadrature component $n_Q(t)$ as follows:

$$n(t) = n_I(t)\cos(2\pi f_c) - n_Q(t)\sin(2\pi f_c)$$

So, we have the following properties

Properties

First: In-phase $n_I(t)$ and quadrature $n_Q(t)$ narrowband noise $n(t)$ have zero mean.

Second: If narrowband noise $n(t)$ is Gaussian, its in-phase and quadrature components are jointly Gaussian.

Third: If narrowband noise is weakly stationary, its in-phase and quadrature are jointly weakly stationary.

Forth: Both in-phase and quadrature noise have same power spectral density, which is related to the power spectral density $S_{NN}(f)$ of the original narrowband noise

Continue Properties

Fifth: In-phase and quadrature components have the same variance as the narrowband noise $n(t)$.

Six: Cross-spectral densities of in-phase & quadrature components of a narrowband noise are pure imaginary:

Seventh: If narrowband noise $n(t)$ is Gaussian with zero mean and $S_N(f)$ power spectral density, that is locally symmetric about the midband frequency $\pm f_c$, the in-phase and quadrature noise are statistically independent.